NIET Journal of Engineering & Technology (NIETJET)

Volume 9, Issue Summer 2021 ISSN: 2229-5828 (Print)

A Literature Review on Web Mining & its Scope in Information Management

Mohammad Shahid

[1]Department of computer science &Engineering, GH Raisoni College of Engineering ,
Nagpur, India.

Abstract: ***Data mining is a critical phase in the information discovery process in databases, and it is a prominent subfield in knowledge management. In the future decades, data mining research will continue to flourish in business and learning organizations. The uses This review research investigates a variety of data mining approaches that have been created to aid in the knowledge management process. The findings are divided into four categories: I knowledge resource; II knowledge categories and/or datasets; III data mining tasks; IV data mining approaches and applications in knowledge management The definition of data mining and data mining functionality are briefly described in the first section of the article. The reasons for knowledge management and the primary knowledge management instruments used in the knowledge management cycle consists of then discussed. Finally, the use of data mining tools between knowledge management pask is described and addressed.***

Keywords: Web Mining, Data Mining, Information management, web mining Applications

Introduction

In the digital era, knowledge is becoming a vital business resource, giving a competitive advantage and prompting knowledge management (KM) activities. Many people have acquired and stored a vast amount of data organizations. They are, which ever, unable to extract important information from data by changing it into valuable and usable knowledge [2]. It's challenging to manage knowledge resources. Many companies are relying on information technology to assist in knowledge creation, sharing, integration, and delivery. Knowledge management is a data-driven process [6]. Data mining is a tool-assisted approach for extracting valuable knowledge from large datasets; it is an essential part of knowledge management. [6].

There are two basic approaches: I sharing shared knowledge about the BI environment among data miners, and (ii) using data mining as a tool to increase human knowledge. As a result, data mining techniques may be able to assist businesses in discovering hidden knowledge inside massive amounts of data.This report employs a literature evaluation of papers from 2007 to 2012 to assess data mining applications in knowledge management as part of data mining research. The rationale for examining research articles during this time period is because data mining has been a hot topic in knowledge management research since 2006 [17], and it serves as a vital connection between business intelligence and knowledge management [26].

We scan the Science Direct database terms "Mining of data" and "management of knowledge" for title keyfield and keywords to filter articles. document search is limited to documents published between 2007 and 2012, and only the document Figure 1 displays the total no of papers produced for this query by year.

The whole content scrutinised to exclude those that are unrelated to data mining's use in knowledge management and do not explain how data mining may be used or aided in knowledge management. These selection criteria are the subject of ten articles.

We study and classify data mining approaches into six categories based on ten papers on data mining applications for knowledge management: categorization, regression, clustering, dependency modelling, deviation detection, and summarization are some of the techniques used in data analysis.

Publisher: Noida Institute of Engineering & Technology,
19, Knowledge Park-II, Institutional Area, Greater Noida (UP), India.

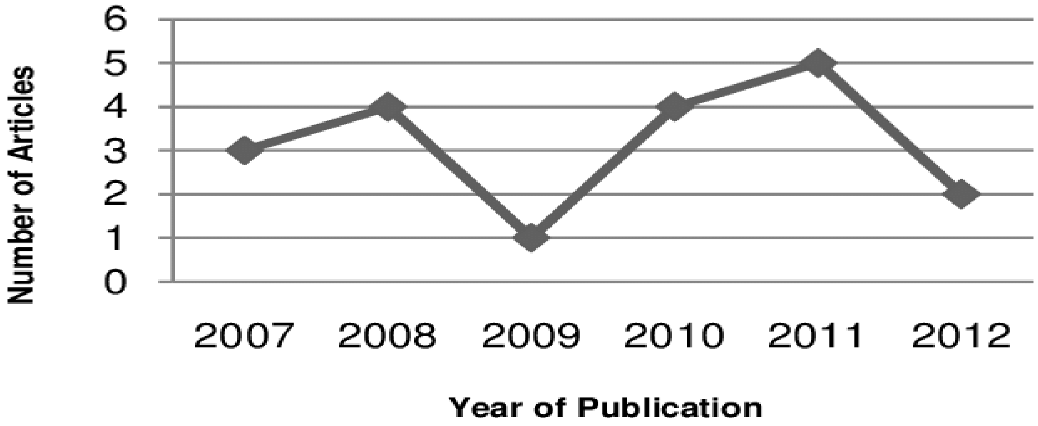

Fig 1: The "Web Mining" or "Information Management"

Web Mining

Content of Web Mining:

Data mining [7] is a crucial phase in the knowledge discovery in databases (KDD) process for generating meaningful patterns or models from data. The phrases KDD and data mining are not interchangeable. The term "knowledge discovery from data" (KDD) refers to the general process of extracting usable information from data. Data mining is the process of extracting usable knowledge from a large amount of data in databases by concentrating on algorithms [7].

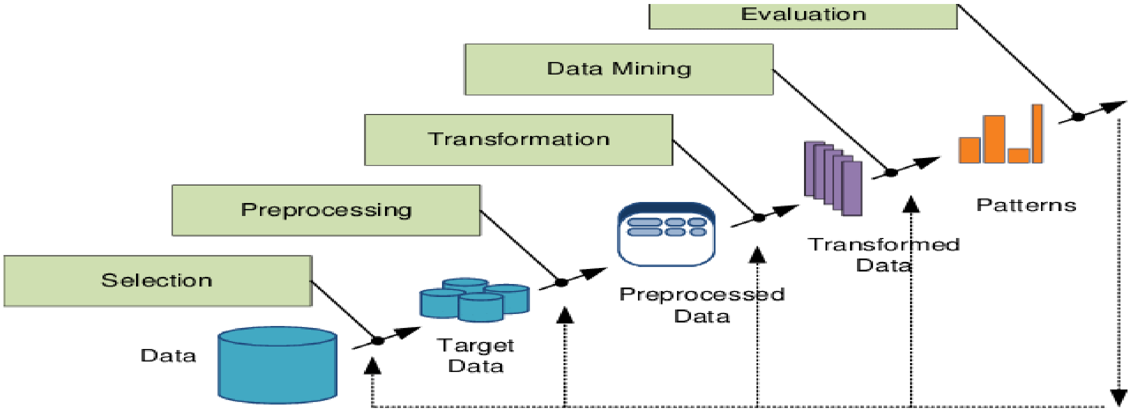

Fig 2: The KDD Process and Web Mining

1. Data selection: Choosing data from the database that is relevant to the analytical activity.

2. Preprocessing: removing noise and inconsistencies from data; merging data from several sources

3. Data transformation: converting data into useful formats for data mining.

4. Data mining: selecting a data mining technique that is suited for the data patterns; extracting data patterns

Publisher: Noida Institute of Engineering & Technology,
19, Knowledge Park-II, Institutional Area, Greater Noida (UP), India.

5. Evaluation/interpretation: converting patterns into knowledge by deleting duplicate or unnecessary patterns; translating beneficial patterns into human-understandable words.

Mining Task:

1. Finding models that examine and categorise a data item into numerous predetermined classifications is known as classification.

2. The mapping of a data item to a real-valued prediction variable is known as regression.

3. Clustering is the process of determining a limited number of categories or clusters to explain a piece of data.

4. Finding a model that reflects substantial interdependencies between variables is known as dependency modelling (Association Rule Learning).

5. Anomaly Detection (Deviation Detection) is the process of identifying the most significant changes in data.

6. Finding a concise explanation for a portion of data is summarization.

Prediction and description are the two main goals of data mining. Prediction (e.g., classification, regression, and anomaly detection) entails utilising some variables in data sets to predict unknown values of other important variables. Clustering, association rule learning, and summarization are examples of methods for detecting human-understandable patterns and trends in data [8].

Information management

Knowledge management may be defined in a variety of ways. In this study, we utilise McInerney's (2002) definition of knowledge management:"Knowledge management (KM) is an endeavor to improve the organization's usable knowledge. Encourage dialogue, provide chances to learn, and promote the sharing of suitable knowledge objects are some of the ways to achieve this. "The interaction component of knowledge management and organizational learning is highlighted in this concept.

The knowledge management process [5] focuses on knowledge flows and the processes of knowledge generation, sharing, and distribution. Information technology can help with each of the knowledge components of capture and generation, sharing and dissemination, and acquisition and application.

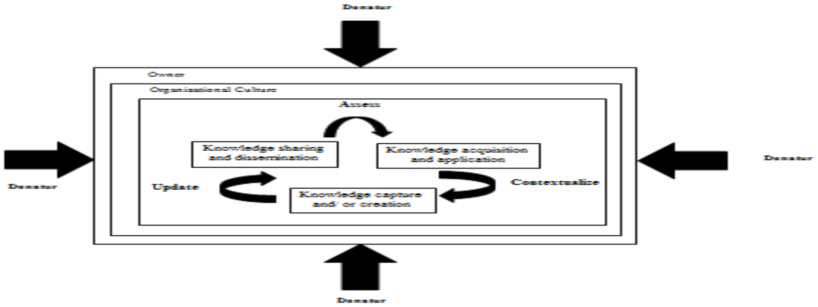
Denatur
Owner
Organizational Culture
Assess
Knowledge sharing and dissemination
Knowledge acquisition and application
Denatur
Denatur
Update
Knowledge capture and/ or creation
Contextualize
Denatur

Fig 3: Integrated KM Cycle by KM Technologies

Because technology plays such an essential part in KM, it is a crucial tool for KM utilisation [1]. For improved knowledge collection, sharing, distribution, and application, KM requires tools that ease communication, collaboration, and content [5].

Knowledge Information

The applications of data mining to organisational knowledge management for effective collecting, storing, retrieving, and transferring information were covered in the reviews of 10 papers. The papers we looked at were sorted into four categories: I knowledge resources; (ii) knowledge kinds and/or datasets; (iii) data mining tasks; and (iv) data mining methodologies and applications utilised in knowledge management.

Information type

The dataset for the Health-Care System domain is made up of three databases: the database for health-care providers, the database for out-patient health-care statistics, and the database for medical status [11]. Another data source was inpatient medical records at a hospital [10].

Customer data and items purchased in the retail realm have been gathered and preserved in databases to see if consumers' buying patterns and behaviour impact product lines and brand expansions [14].

Conclusion

Knowledge is a valuable resource in every firm. The management of knowledge resources has become a major development priority. Discovering meaningful information is also an important strategy for management and decision-making. Because data mining is such an important aspect of knowledge management, this research looked for 10 publications about data mining applications in knowledge management that were published between 2007 and 2012. This paper seeks to provide a study summary on data mining's applicability in the KM technologies arena. The findings in this research are based on the following

1. Based on the publishing rates, study on the use of data mining in knowledge management will grow in the future, covering a wide range of interests.
2. The use of a problem-oriented domain to KM application development tends to promote expert decision making.

References

[1] An, X. & Wang, W. (2010). Knowledge management technologies and applications: A literature review. IEEE, 138-141. doi:10.1109/ICAMS.2010.5553046

[2] Berson, A., Smith, S.J. &Thearling, K. (1999). Building Data Mining Applications for CRM. New York: McGraw-Hill.

[3] Cantú, F.J. & Ceballos, H.G. (2010). A multiagent knowledge and information network approach for managing research assets. Expert Systems with Applications, 37(7), 5272-5284. doi:10.1016/j.eswa.2010.01.012

[4] Cheng, H., Lu, Y. & Sheu, C. (2009). An ontology-based business intelligence application in a financial knowledge management system.Expert Systems with Applications, 36, 3614–3622. Doi:10.1016/j.eswa.2008.02.047 International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.2, No.5, September 2012 24

[5] Dalkir, K. (2005). Knowledge Management in Theory and Practice. Boston: Butterworth-Heinemann.

[6] Dawei, J. (2011). The Application of Date Mining in Knowledge Management.2011 International Conference on Management of e-Commerce and e-Government, IEEE Computer Society, 7-9. doi: 10.1109/ICMeCG.2011.58 [7] Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases.AI Magazine, 17(3), 37-54.

[7] Gorunescu, F. (2011). Data Mining: Concepts, Models, and Techniques. India: Springer.

[8] Han, J. &Kamber, M. (2012). Data Mining: Concepts and Techniques. 3rd.ed. Boston: Morgan Kaufmann Publishers.

[9] Hwang, H.G., Chang, I.C., Chen, F.J. & Wu, S.Y. (2008). Investigation of the application of KMS for diseases classifications: A study in a Taiwanese hospital. Expert Systems with Applications, 34(1), 725-733. doi:10.1016/j.eswa.2006.10.018

[10] Lavrac, N., Bohanec, M., Pur, A., Cestnik, B., Debeljak, M. &Kobler, A. (2007).Data mining and visualization for decision support and modeling of public health-care resources.Journal of Biomedical Informatics, 40, 438-447. doi:10.1016/j.jbi.2006.10.003

[11] Li, X., Zhu, Z. & Pan, X. (2010). Knowledge cultivating for intelligent decision making in small & middle businesses.Procedia Computer Science, 1(1), 2479-2488. doi:10.1016/j.procs.2010.04.280

[12] Li, Y., Kramer, M.R., Beulens, A.J.M., Van Der Vorst, J.G.A.J. (2010). A framework for early warning and proactive control systems in food supply chain networks. Computers in Industry, 61, 852–862. Doi:101.016/j.compind.2010.07.010

[13] Liao, S.H., Chen, C.M., Wu, C.H. (2008). Mining customer knowledge for product line and brand extension in retailing.Expert Systems with Applications, 34(3), 1763-1776. doi:10.1016/j.eswa.2007.01.036

[14] Liao, S. (2003). Knowledge management technologies and applications-literature review from 1995 to 2002. Expert Systems with Applications, 25, 155-164. doi:10.1016/S0957-4174(03)00043-5

[15] Liu, D.R. & Lai, C.H. (2011). Mining group-based knowledge flows for sharing task knowledge. Decision Support Systems,50(2), 370-386. doi:10.1016/j.dss.2010.09.004 .

[16] Ayushi Singhal , Amrendra Singh Yadav ,Mayank Deep Khare, Rahul Kumar Sharma; 'Integrated MANET Mutual Authen-tication System -A Review',Volume No.4,Issue No.2,2016,PP.044-052,ISSN :2229-5828

[17] Jitendra Kumar Saroj, V.K.Pandey; 'Micro -strip Slot Antenna with High Efficiency',Volume No.4,Issue No.2,2016,PP.053-056,ISSN :2229-5828

[18] Pradip Kumar Yadava, Surya Deo Choudhary; 'Design & Performance Analysis of CPW Fed Square slot Antenna and Horizontal H-Shaped Stub Antenna',Volume No.4,Issue No.2,2016,PP.057-061,ISSN :2229-5828

[19] Deepak Kumar, Harendra Singhal, Somesh Kumar; 'Simulation of Steady State and Dynamic Response of Multi -effect Evaporators in Paper Industry',Volume No.4,Issue No.2,2016,PP.062-070,ISSN :2229-5828

[20] Rahul Yadav, Sanjay Gairola; 'Zigzag Connected Autotran-sformer Based 12,24 &36-Pulse Rectifiers',Volume No.4,Issue No.2,2016,PP.071-076,ISSN :2229-5828

[21] P.Kumar,R.K.Rajuvanshi,R.sharma, P.Yadav& P.Chaturvedi; 'Energy Auditing and Management :A Case Study to Improve Energy Efficiency and Setting Benchmarking',Volume No.1,Issue No.1,2012,PP.001-007,ISSN :2229-5828